Time Series and their Applications

Homework Exercise 5

by

Dico de Gier (SNR: 2058017)

Richard Brijder (SNR: 2119548)

Freek Verstraaten (SNR: 2059593)

Date: May 31, 2024

Question 2

We show that $P_T X_{t+h}$, $h \geq 1$, based on $\{W, X_1, \ldots, X_T\}$ is the same as the best linear forecast of X_{t+h} based on $\boldsymbol{X_T} = \{X_1, \ldots, X_T\}$. Otherwise stated, an observation of W does not impact the best linear forecast. Notice that the best linear predictor based on $\{W, X_1, \ldots, X_T\}$ minimizes by definition

$$\mathbb{E}\left[(X_{t+h} - g(\boldsymbol{X_T}, W))^2\right]$$

over all functions g that are a linear combination of $\boldsymbol{X_T}$ and W. Therefore $g(\boldsymbol{X_T}, W)$ has the form

$$g(\boldsymbol{X_T}, W) = a_0 + a_1 X_1 + \cdots + a_T X_T + \lambda W.$$

We prove $\lambda = 0$, because in this case an observation of W does not impact the best linear forecast.

Define S as

$$S(a_0, \ldots, a_T, \lambda) := \mathbb{E}\left[(X_{t+h} - a_0 - a_1 X_1 - \cdots - a_T X_T - \lambda W)^2\right].$$

The first-order condition with respect to λ for the minimization of S is given by

$$\begin{aligned}\frac{\partial S(a_0, \ldots, a_T, \lambda)}{\partial \lambda} &= -2\mathbb{E}\left[(X_{t+h} - a_0 - a_1 X_1 - \cdots - a_T X_T - \lambda W)\, W\right] \\ &= -2\left(\mathbb{E}[X_{t+h} W] - a_0 \mathbb{E}[W] - a_1 \mathbb{E}[X_1 W] - \cdots - \mathbb{E}[X_T W] - \lambda \mathbb{E}[W^2]\right)\end{aligned}$$

where $\mathbb{E}[X_t W] = 0$ for all t. Hence

$$\frac{\partial S(a_0, \ldots, a_T, \lambda)}{\partial \lambda} = 2a_0 \mathbb{E}[W] + 2\lambda \mathbb{E}[W^2]$$

which is zero if and only if $\lambda = -a_0 \frac{\mathbb{E}[W]}{\mathbb{E}[W^2]}$. (This is well-defined, since $\mathbb{E}[W^2] < \inf$.) Notice we know this λ minimizes the function S, because

$$\frac{\partial^2 S(a_0, \ldots, a_T, \lambda)}{\partial \lambda^2} = 2\mathbb{E}[W^2] > 0.$$

Notice on the other hand, that the minimization of S with respect to a_0 implies the condition

$$\begin{aligned}\frac{\partial S(a_0, \ldots, a_T, \lambda)}{\partial a_0} &= -2\mathbb{E}\left[(X_{t+h} - a_0 - a_1 X_1 - \cdots - a_T X_T - \lambda W)\right] \\ &= -2\left(\mathbb{E}[X_{t+h}] - a_0 - a_1 \mathbb{E}[X_1] - \cdots - a_T \mathbb{E}[X_T] - \lambda \mathbb{E}[W]\right)\end{aligned}$$

which equals, because $\{X_t\}$ has mean zero,

$$\frac{\partial S(a_0, \ldots, a_T, \lambda)}{\partial a_0} = 2a_0 + 2\lambda \mathbb{E}[W]$$

which is zero if and only if $a_0 = -\lambda \mathbb{E}[W]$. Notice this a_0 minimizes S, since

$$\frac{\partial^2 S(a_0, \ldots, a_T, \lambda)}{\partial a_0^2} = 2 > 0.$$

From this we deduce that the λ for which S is minimized equals

$$\lambda = a_0 \frac{\mathbb{E}[W]}{\mathbb{E}[W^2]} = -\lambda \frac{(\mathbb{E}[W])^2}{\mathbb{E}[W^2]}.$$

Since both $(\mathbb{E}[W])^2$ and $\mathbb{E}[W^2]$ are non-negative, this implies that $\lambda = 0$. As a result, we see that W does not impact the best linear forecast and hence the linear forecast based on $\{X_1, \ldots, X_T\}$ is the same as $P_T X_{t+h}$, $h \geq 1$, based on $\{W, X_1, \ldots, X_T\}$.

Question 3

$$
\begin{aligned}
X_t - \mu &= \sum_{j=1}^{p} \phi_j (X_{t-j} - \mu) + Z_t \\
X_t - \mu - X_{t-1} &= \sum_{j=1}^{p} \phi_j (X_{t-j} - \mu) + Z_t - X_{t-1} \\
\Delta X_t &= -\sum_{j=1}^{p} \phi_j \mu + \sum_{j=1}^{p} \phi_j X_{t-j} + Z_t - X_{t-1} + \mu \\
\Delta X_t &= \mu(1 - \sum_{j=1}^{p} \phi_j) + \sum_{j=1}^{p} \phi_j X_{t-j} + Z_t - X_{t-1} \\
\Delta X_t &= \gamma_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + Z_t - X_{t-1} \\
\Delta X_t &= \gamma_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + Z_t - X_{t-1} + \phi_2 X_{t-1} - \phi_2 X_{t-1} \\
&\quad + \phi_3 X_{t-1} - \phi_3 X_{t-1} + ... + \phi_p X_{t-1} - \phi_p X_{t-1} \\
\Delta X_t &= \gamma_0 + X_{t-1}(\phi_1 + \phi_2 + ... + \phi_p) - X_{t-1}(\phi_2 + \phi_3 + ... + \phi_p) + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + Z_t \\
\Delta X_t &= \gamma_0 + \gamma_1 X_{t-1} - X_{t-1}(\phi_2 + \phi_3 + ... + \phi_p) + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + Z_t
\end{aligned}
$$

For any $j \geq 3$, we have:

$$
\gamma_j - \gamma_{j-1} = -\sum_{i=j}^{p} \phi_i + \sum_{i=j-1}^{p} \phi_i = -(\phi_j + \phi_{j+1} + ... + \phi_p) + (\phi_{j-1} + \phi_j + ... + \phi_p) = \phi_{j-1}
$$

Then we find that:

$$
\begin{aligned}
\sum_{j=1}^{p-1} \gamma_{j+1} \Delta X_{t-j} &= \gamma_2 X_{t-1} - \gamma_2 X_{t-2} + \gamma_3 X_{t-2} - \gamma_3 X_{t-3} + ... + \gamma_p X_{t-p+1} - \gamma_2 X_{t-p} \\
&= \gamma_2 X_{t-1} + X_{t-2}(\gamma_3 - \gamma_2) + X_{t-3}(\gamma_4 - \gamma_3) + ... + X_{t-p+1}(\gamma_p - \gamma_{p-1}) - \gamma_p X_{t-p} \\
&= -X_{t-1} \sum_{i=2}^{p} \phi_i + \phi_2 X_{t-2} + ... + \phi_{p-1} X_{t-p+1} - -X_{t-p} \sum_{i=p}^{p} \phi_i \\
&= -X_{t-1}(\phi_2 + \phi_3 + ... + \phi_p) + \phi_2 X_{t-2} + ... + \phi_p X_{t-p}
\end{aligned}
$$

We can now plug this into the first equation we derived, obtaining the following result:

$$
\Delta X_t = \gamma_0 + \gamma_1 X_{t-1} + \sum_{j=1}^{p-1} \gamma_{j+1} \Delta X_{t-j} + Z_t
$$

Question 4

A

For lag 0, we know

$$
\gamma_X(0) = \int_{-\frac{1}{2}}^{\frac{1}{2}} f_X(\omega) \mathrm{d}\omega = \int_{0}^{\frac{1}{2}} 2 f_X(\omega) \mathrm{d}\omega
$$

because $f_X(\omega) = f_X(-\omega)$ for $\omega \in [-\frac{1}{2}, 0]$. Moreover,

$$\gamma_X(0) = \int_0^{\frac{1}{2}} 2f_X(\omega)\mathrm{d}\omega = \int_0^{\frac{1}{2}} 200 \cdot 1_{\omega\in(\frac{1}{3}-0.01,\frac{1}{3}+0.01)}\mathrm{d}\omega = \int_{\frac{1}{3}-0.01}^{\frac{1}{3}+0.01} 200\mathrm{d}\omega$$
$$= 200(\frac{1}{3} + 0.01 - (\frac{1}{3} - 0.01)) = 200 * 0.02 = 4.$$

For lag $h = 1$, we know

$$\begin{aligned}
\gamma_X(1) &= \int_0^{\frac{1}{2}} 2f_X(\omega)e^{2\pi i\omega}\mathrm{d}\omega = \int_{\frac{1}{3}-0.01}^{\frac{1}{3}+0.01} 200e^{2\pi i\omega}\mathrm{d}\omega \\
&= 200\left[\frac{e^{2\pi i\omega}}{2\pi i}\right]_{\frac{1}{3}-0.01}^{\frac{1}{3}+0.01} \\
&= \frac{100}{\pi i}\left(e^{\frac{2\pi i}{3}+0.02\pi i} - e^{\frac{2\pi i}{3}-0.02\pi i}\right) \\
&= \frac{100}{\pi i}\left(\cos(\frac{2\pi}{3}+0.02\pi) - \cos(\frac{2\pi}{3}-0.02\pi) + i\left(\sin(\frac{2\pi}{3}+0.02\pi) - \sin(\frac{2\pi}{3}-0.02\pi)\right)\right)
\end{aligned}$$

where we used $e^{i\omega} = \cos(\omega) + i\sin(\omega)$.

B

Define Ψ as $\Psi(L) = (1 - L^{12})$. Then $Y_t = \Psi(L)X_t$ and the spectrum is given by

$$\begin{aligned}
f_Y(\omega) &= \left|\hat{\Psi}(\omega)\right|^2 f_X(\omega) \\
&= \Psi(e^{-2\pi i\omega})\Psi(e^{2\pi i\omega})f_X(\omega) \\
&= \left(1 - e^{-24\pi iw}\right)\left(1 - e^{24\pi iw}\right)f_X(\omega) \\
&= \left(1 - e^{-24\pi iw} - e^{24\pi iw} + 1\right)f_X(\omega) \\
&= (2 - 2\cos(24\pi\omega))\, f_X(\omega) \\
&= (200 - 200\cos(24\pi\omega))\, 1_{\omega\in(\frac{1}{3}-0.01,\frac{1}{3}+0.01)}
\end{aligned}$$

where we used $(e^{i\omega} + e^{-i\omega})/2 = \cos(\omega)$.

C

The variance of Y_t is given by

$$\begin{aligned}
\gamma_Y(0) &= \int_{-\frac{1}{2}}^{\frac{1}{2}} f_Y(\omega)\mathrm{d}\omega \\
&= 200\int_{\frac{1}{3}-0.01}^{\frac{1}{3}+0.01} 2 - 2\cos(24\pi\omega)\mathrm{d}\omega \\
&= 200\left[2\omega - \frac{\sin(24\pi\omega)}{12\pi}\right]_{\frac{1}{3}-0.01}^{\frac{1}{3}+0.01} \\
&= 8 - \frac{50}{3\pi}\left(\sin(8.24\pi) - \sin(7.76\pi)\right).
\end{aligned}$$

D

The transfer function $\hat{\Psi}$ is given by

$$\hat{\Psi}(\omega) = \Psi(e^{-2\pi i\omega}) = \left(1 - e^{-24\pi i\omega}\right)$$

with the corresponding plot of the real part:

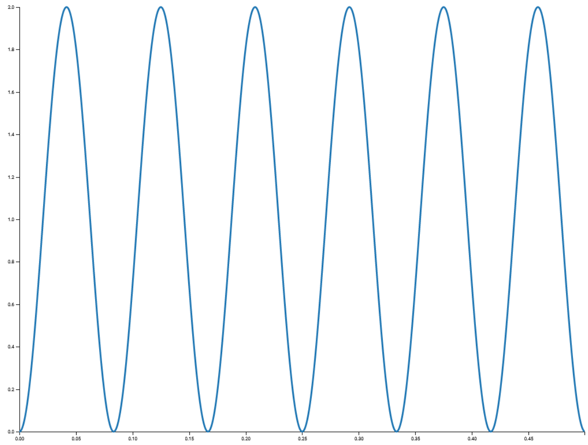

Figure 1: Transfer function for $\omega \in [0, \frac{1}{2}]$

Notice this function removes the seasonality for each month of the process, which can be seen from the fact that the transfer function is zero for all multiples of $\omega_i = 1/12$. It eliminates all frequencies that are multiples of $1/12$, including zero.

Question 5

We verify the convolution property, using the inverse DFT as given in the exercise,

$$\begin{aligned}
\sum_{s=1}^{T} a_s X_{t-s} &= a_1 X_{t-1} + \cdots + a_T X_{t-T} \\
&= \frac{1}{\sqrt{T}} \left(a_1 \sum_{i=0}^{T-1} d_X(\omega_i) e^{2\pi\tau\omega_i(t-1)} + \cdots + a_T \sum_{i=0}^{T-1} d_X(\omega_i) e^{2\pi\tau\omega_i(t-T)} \right) \\
&= \frac{1}{\sqrt{T}} \left(\sum_{s=1}^{T} a_s \sum_{i=0}^{T-1} d_X(\omega_i) e^{2\pi\tau\omega_i(t-s)} \right)
\end{aligned}$$

where τ is the imaginary unit. Notice that a_s does not depend on i for $s = 1, \dots, T$ and $i = 0, \dots, T-1$. Hence we can interchange summations,

$$\begin{aligned}
\sum_{s=1}^{T} a_s X_{t-s} &= \frac{1}{\sqrt{T}} \left(\sum_{i=0}^{T-1} d_X(\omega_i) \sum_{s=1}^{T} a_s e^{2\pi\tau\omega_i(t-s)} \right) \\
&= \frac{1}{\sqrt{T}} \left(\sum_{i=0}^{T-1} d_X(\omega_i) e^{2\pi\tau\omega_i t} \sum_{s=1}^{T} a_s e^{-2\pi\tau\omega_i s} \right) \\
&= \sum_{i=0}^{T-1} d_X(\omega_i) e^{2\pi\tau\omega_i t} \frac{1}{\sqrt{T}} \sum_{s=1}^{T} a_s e^{-2\pi\tau\omega_i s} \\
&= \sum_{i=0}^{T-1} d_X(\omega_i) e^{2\pi\tau\omega_i t} d_A(\omega_i) \\
&= \sum_{i=0}^{T-1} d_A(\omega_i) d_X(\omega_i) e^{2\pi\tau\omega_i t}
\end{aligned}$$

where d_A and d_X are the DFTs of a_t and X_t respectively.

Empirical Exercise Part I

(1)

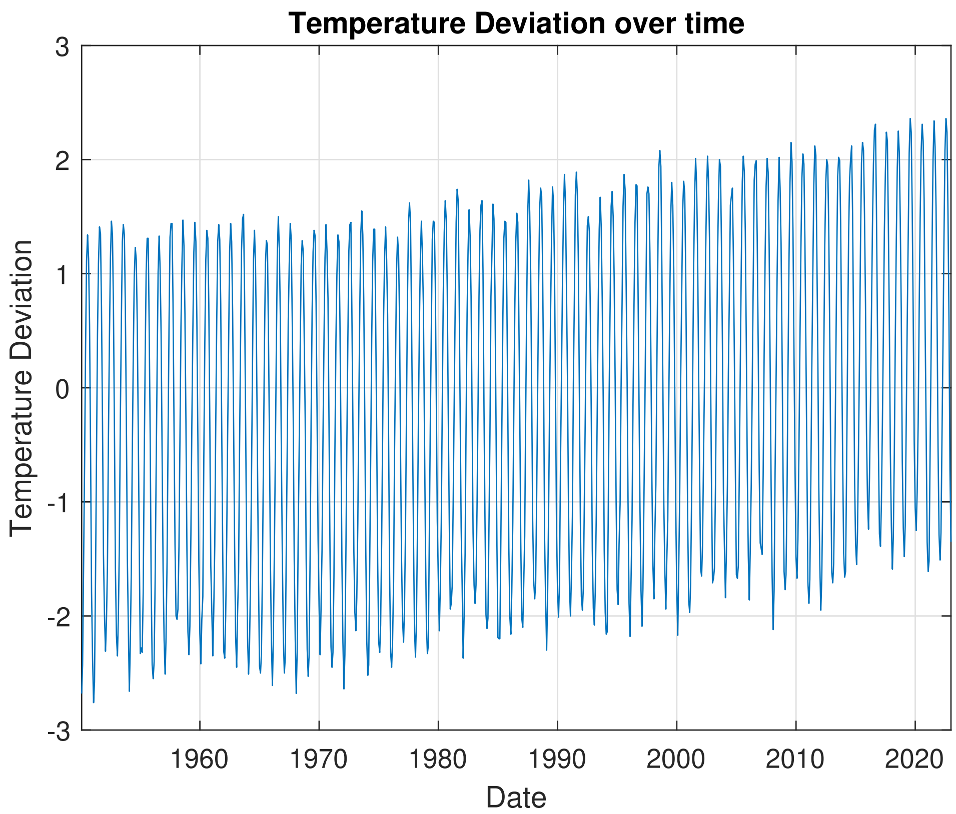

Figure 2: Plot of data against time

From Figure 3, we clearly see that temperature deviations are lower in the winter months and much higher in the summer months. Therefore, the data exhibits a seasonal pattern

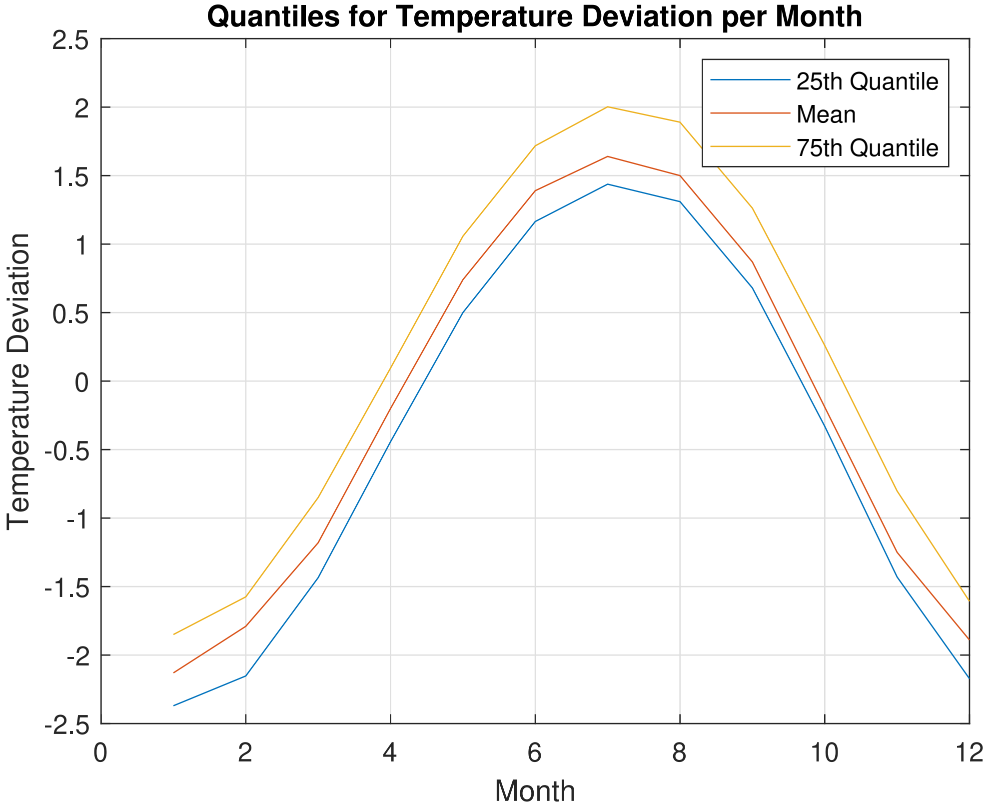

Figure 3: Quantiles per month

(2)

We wish to test the null hypothesis that there is a unit root present against the alternative hypothesis that there is no unit root present. After applying seasonal differencing and performing the ADF test, we arrive at a p-value of 0.0001, meaning that we reject the null hypothesis for any reasonable confidence level. Therefore, there is no unit root present.

In both Figures 4 and 5 we see that there are several lag values for which either the ACF or the PACF lies outside the confidence bands. Therefore, this is evidence that seasonal lags are present in the data.

(3)

As we were initially unsure how to decide on the best model, we asked at the end of the lecture and were told that it is okay to try several values for the parameters and then decide which are "the best". For each choice, we applied $d = 0$ and $D = 1$. The full results can be found in the pdf of the modeler app, but the main results are summarized in Table 1

From Table 1 we conclude that there are two choices based on the lowest AIC. If we then also take into account the number of significant parameters for each of these two models, we conclude that the best model has $p = 4,\ q = 4,\ P = 2$ and $Q = 2$. The results for this model are shown in Table 2

For diagnostic checking, we first determined the sample autocorrelation function, which is shown in Figure 6. This figure shows that all ACFs clearly lie within the confidence bounds. To formally test this, we also applied the Ljung-Box test from the modeler. This resulted in a p-value of 0.9955, hence we clearly do not reject the null hypothesis that the autocorrelations are jointly 0.

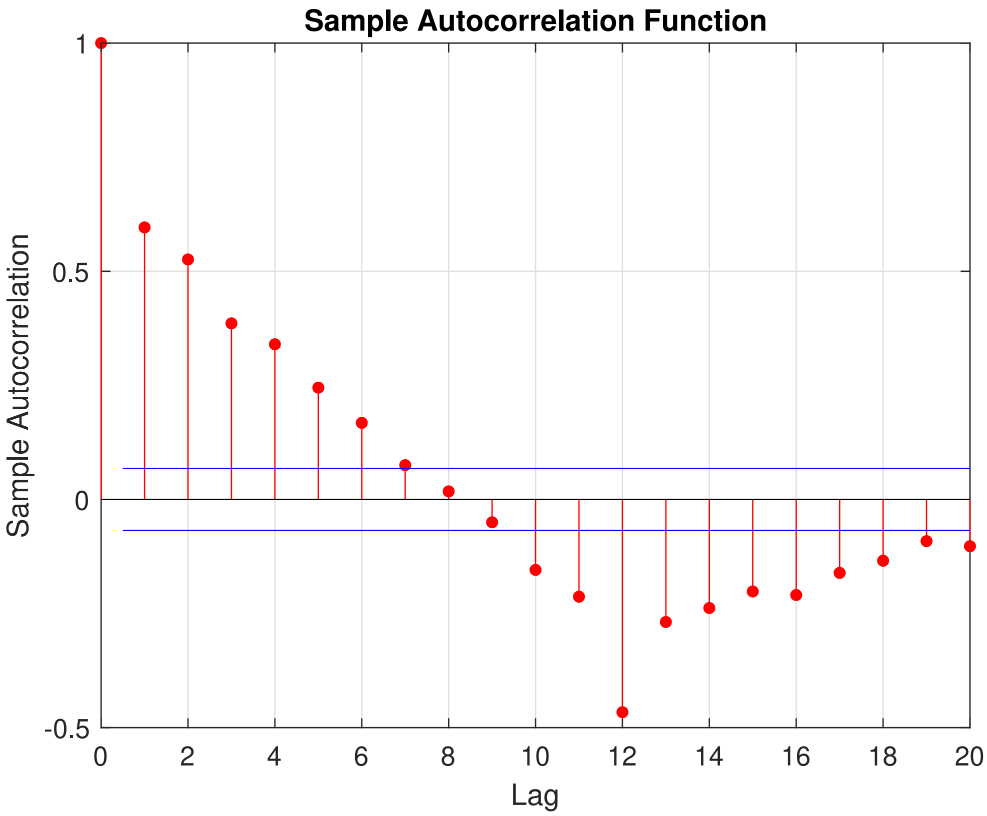

Figure 4: ACF of transformed series

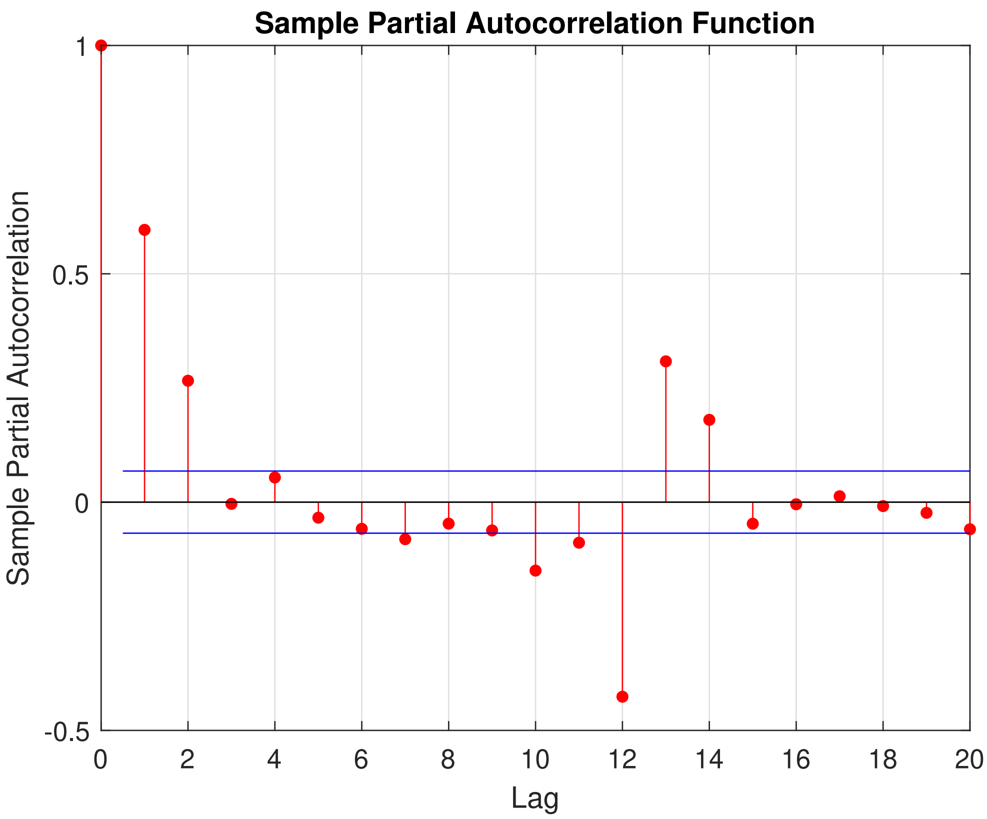

Figure 5: PACF of transformed series

p	q	P	Q	AIC	number of significant parameters [1]
4	4	1	1	-1405	3
3	3	1	1	-1403	4
2	2	1	1	-1397	4
1	1	1	1	-1356	3
4	4	2	2	-1405	7
3	3	2	2	-1399	3
2	2	2	2	-1386	5
1	1	2	2	-1361	3

Table 1: Summary of the results from the modeler app

	Value	Standard Error
Constant	0.0000	0.0000
AR1	0.6115	0.2670
AR2	-0.2587	0.1580
AR3	0.0836	0.1646
AR4	0.2350	0.0900
SAR12	-0.3613	0.1540
SAR24	-0.0142	0.0377
MA1	-1.1534	0.2659
MA2	0.6092	0.2795
MA3	-0.3958	0.1710
MA4	-0.0600	0.1255
SMA12	-0.5863	0.1594
SMA24	-0.2552	0.1402
Variance	0.0114	0.0005

Table 2: Results for $p = 4,\ q = 4,\ P = 2$ and $Q = 2$

(4)

h	Forecast	MSPE
1	-1.5685	0.0114
2	-1.2664	0.0138
3	-0.4914	0.0164
4	0.4426	0.0175
5	1.3600	0.0189
6	2.0380	0.0197
7	2.3209	0.0203
8	2.2198	0.0206
9	1.6116	0.0209
10	0.6488	0.0210
11	-0.4430	0.0212
12	-1.1775	0.0213

Table 3: Forecasts and MSPEs

Figure 7 shows that the forecasts and the realized values are actually quite close to each other. This suggests that the chosen model is quite powerful when it comes to predicting future values of the series.

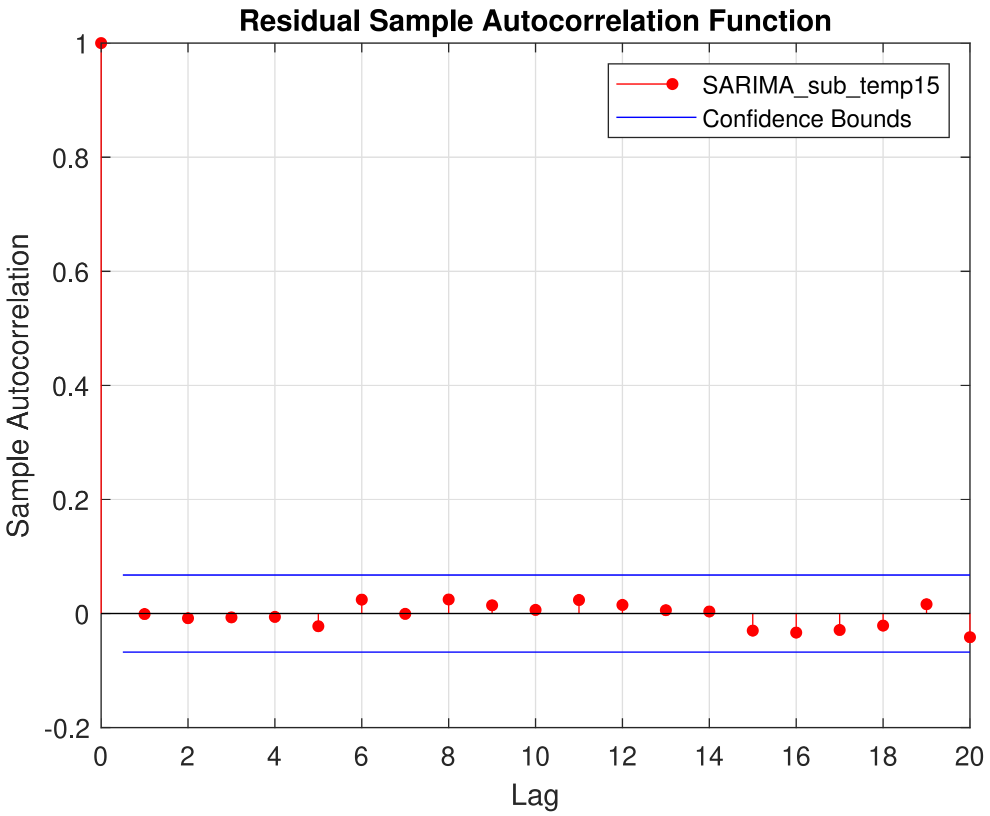

Figure 6: Sample Autocorrelation Function

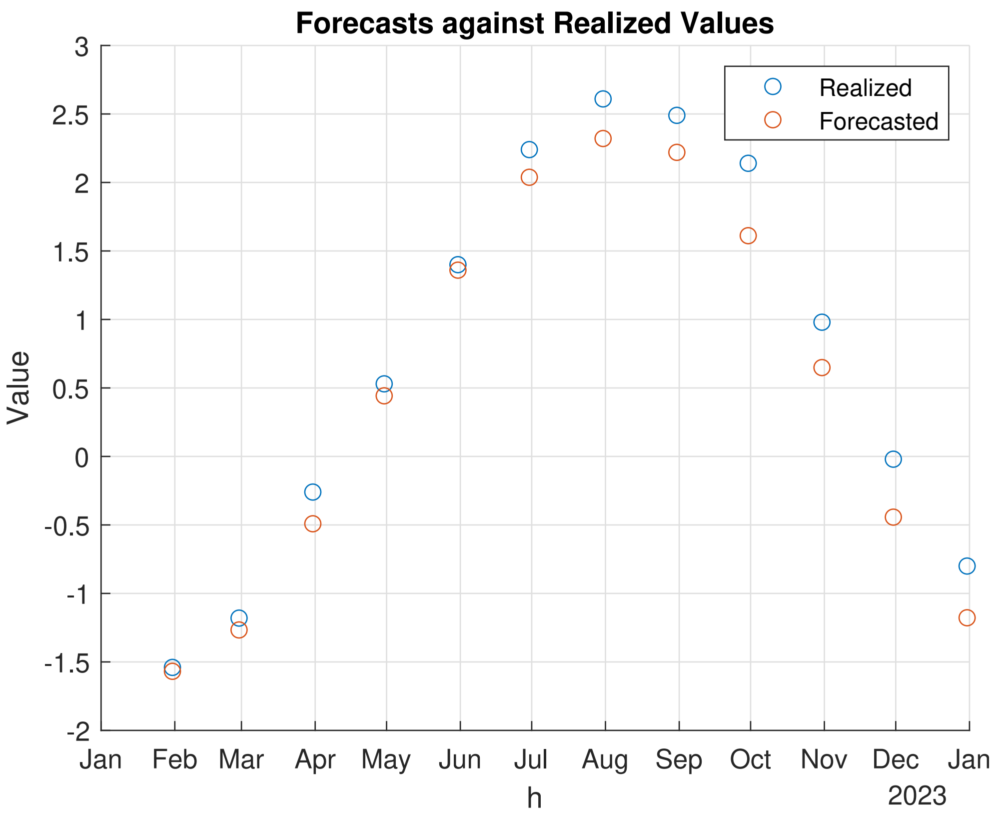

Figure 7: Forecasted against Realized values

Empirical Exerise Part II

(1)

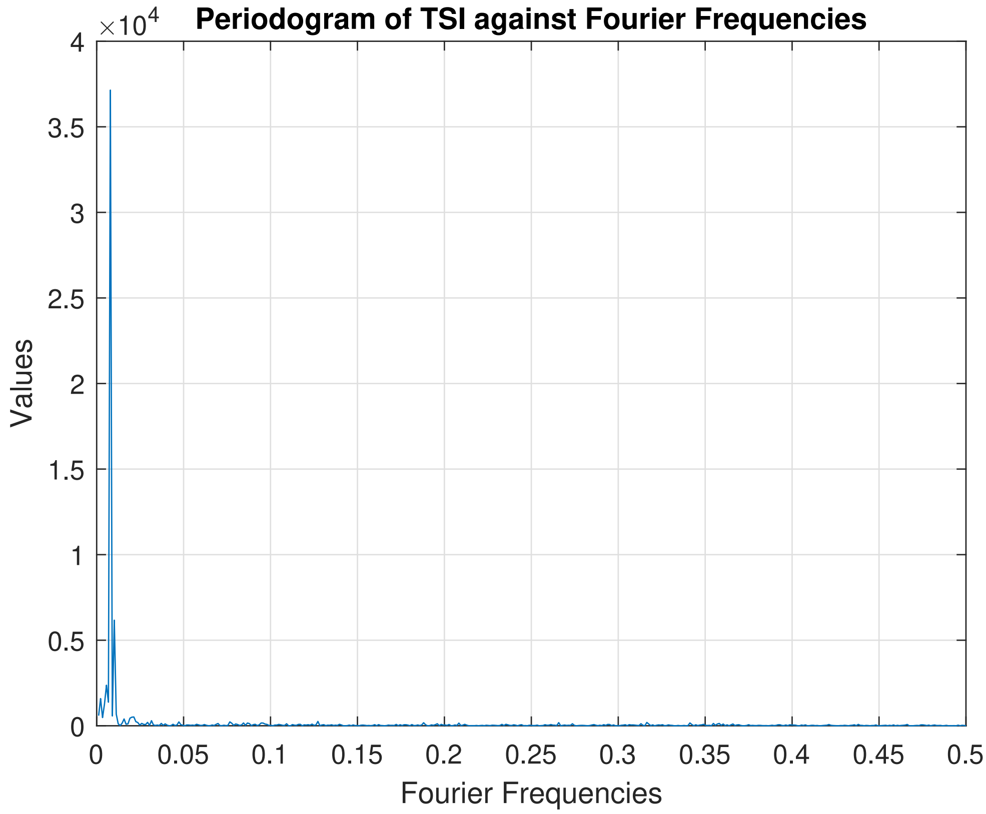

Figure 8: Periodogram of TSI against Fourier frequencies

Since Figure 8 contains one big spike which makes the rest of the plot rather flat, we again show the plot but now with logs on the y-axis in Figure 9.

We see that there is a large spike around the Fourier frequency 0.00788. However, we were unable to find a relation between the most pronounced cycle and the wiki article.

(2)

Initially, we created Figure 10, but as this picture was quite uninformative, we again took logs for the y-axis which resulted in Figure 11

(3)

Initially, we again created a plot without logs (Figure 12) which was again not very informative. A more informative plot with logs on the y-axis is shown in Figure 13.

From Figure 13, we conclude that the model does *not* adequately capture the features of the data, as the implied spectral density and the discrete spectral average estimate are quite far from each other for all frequencies.

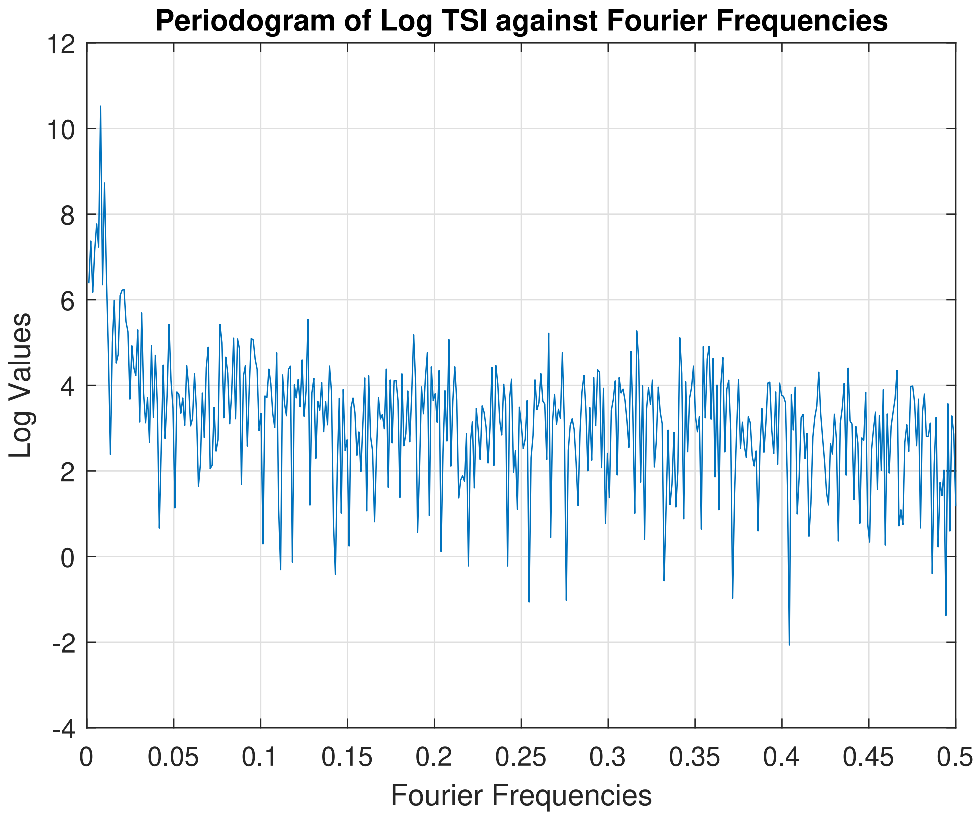

Figure 9: Periodogram of log TSI against Fourier frequencies

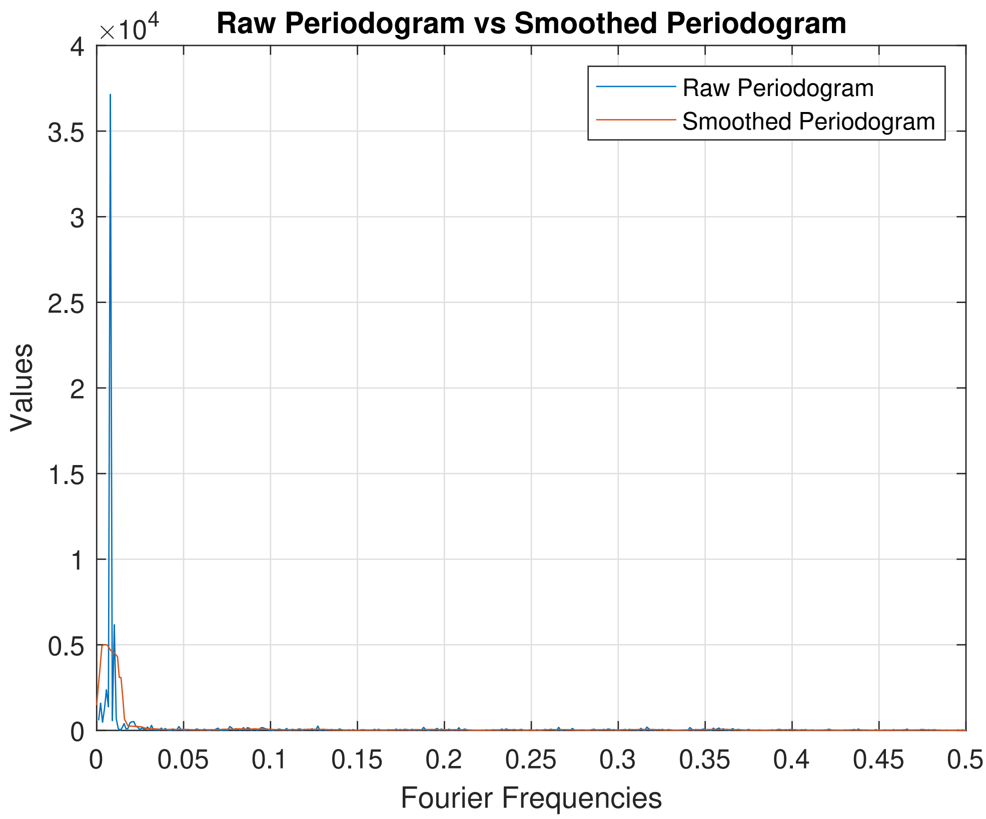

Figure 10: Raw periodogram against smoothed periodogram

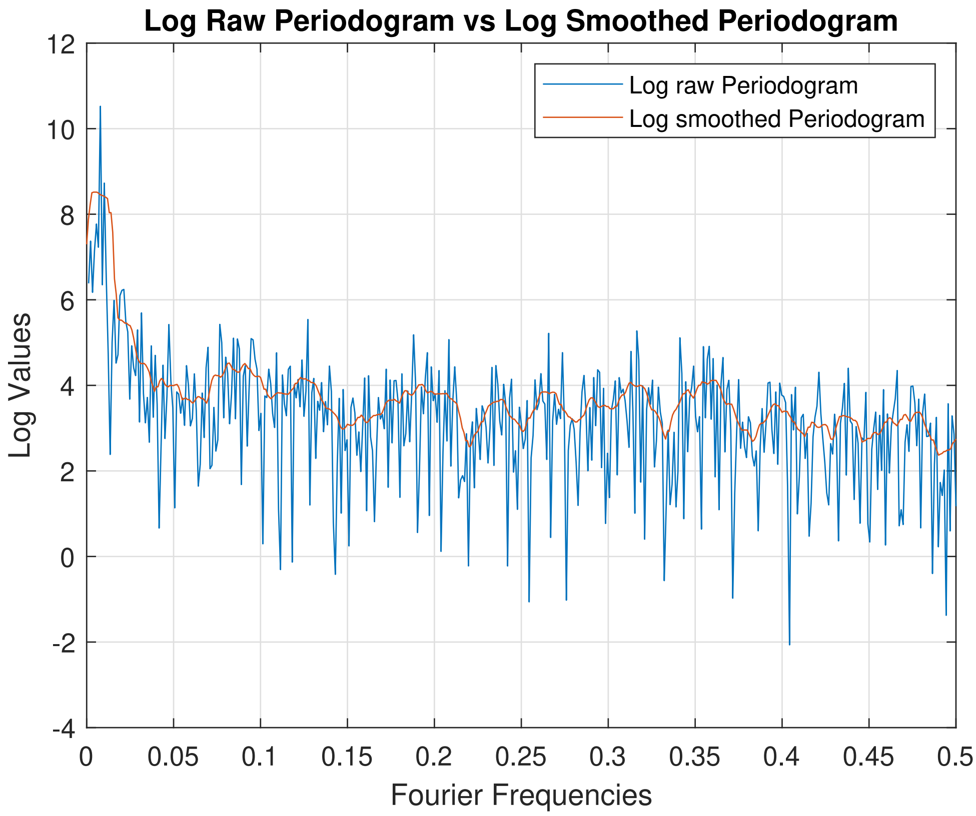

Figure 11: Log raw periodogram against log smoothed periodogram

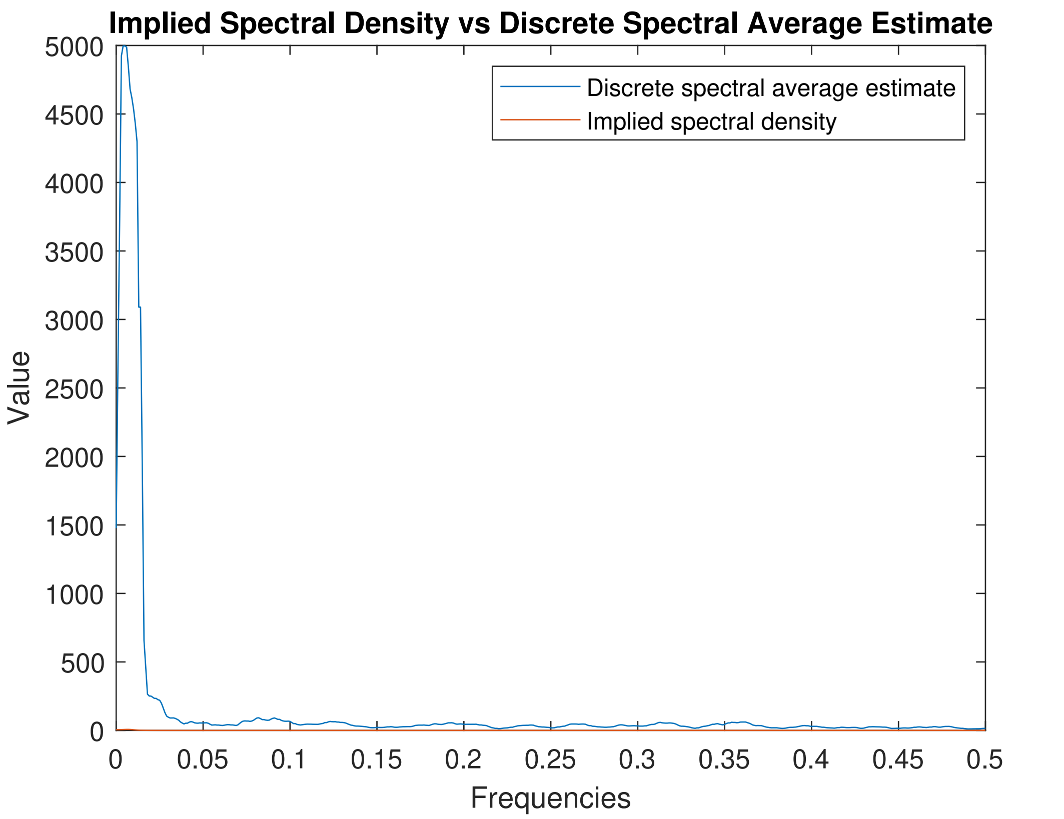

Figure 12: Implied spectral density against discrete spectral average estimate

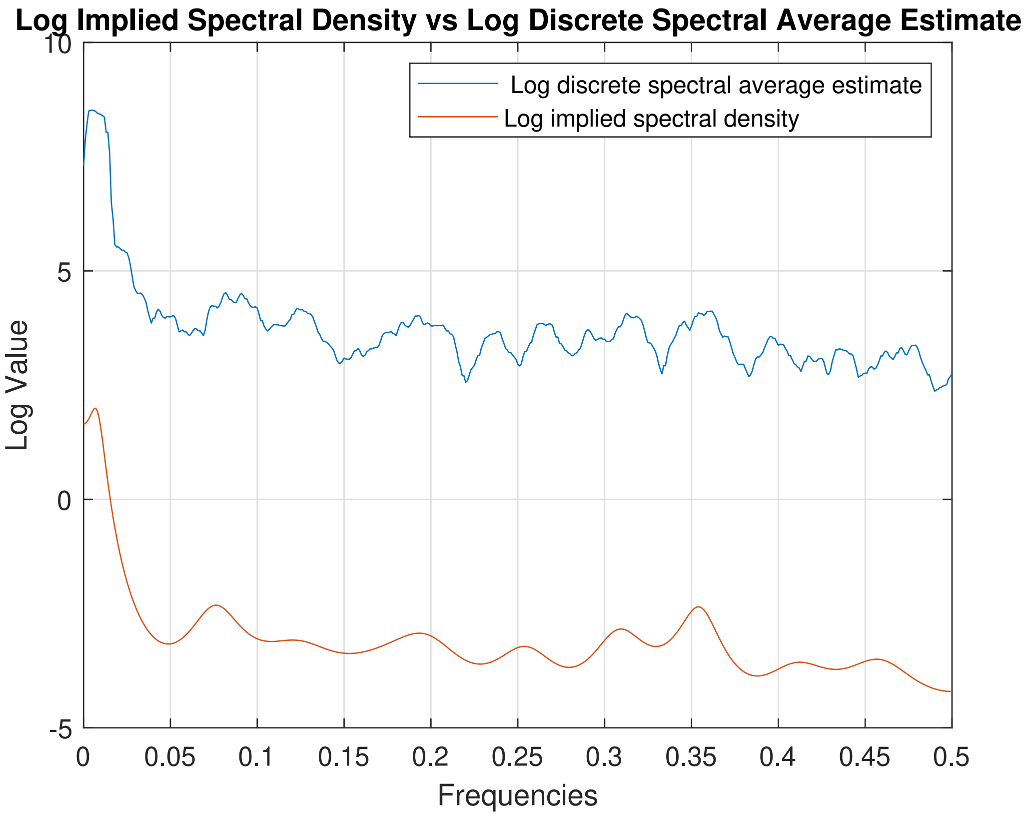

Figure 13: Log implied spectral density against log discrete spectral average estimate